

ドイツ語基本語彙リストの比較

大 藺 正 彦

『ドイツ文学論集』第47号別刷

日本独文学会中国四国支部

2014（平成26）年10月

ドイツ語基本語彙リストの比較

大藪 正彦

1. はじめに

筆者は、頻度に基づくドイツ語基本語彙リスト（5,000語）の作成について報告した際（大藪 2014）、日本におけるドイツ語教育への応用という観点から、今後検討すべき課題として次の2点を挙げた。一つは高頻度の語と学習上優先すべき語のずれの問題、もう一つは母語話者にとっての頻度と日本人にとっての頻度のずれの問題である。¹⁾ これらの問題の実態を把握するため、本稿では、(1) 頻度語彙リストと教育語彙リストの比較、(2) 頻度語彙リストと独自調査による語彙リストの比較を行う。これらの調査を通し、日本人学習者にとってのドイツ語基本語彙とは何かについて考えてみたい。²⁾

1.1 「基本語彙」について

30万から50万とも言われるドイツ語の語彙³⁾の中から、学習者はまずどのような語を学ばよいか。語彙学習を見通しのきくものにするための一つの方法として、いわゆる「基本語彙」(Grundwortschatz)の策定がある。特定の基準によって語数を縮小しようというものである。その基準はさまざまであるが、ここで本稿の考え方を説明しておく。

まず筆者は、「日本におけるドイツ語学習またはドイツ語教育・研究において参照できる、頻度に基づく基本語彙リスト」の作成を構想している。その目標を念頭に置いた上で、千野（1986）による下の図1から出発する。縦軸に頻度、横軸にジャンル——使用域 (Register) と言い換えてもよいだろう——の広さを設定して、A～Dの4区分を立てたものである。⁴⁾ 現在のところ、この線引きは作業上の仮説と考えるべきものであり、具体的にどれく

- 1) 便宜上、本稿では単に日本人という言い方を使うが、言うまでもなく国籍が問題なのではない。想定している典型的なケースは、日本に暮らす（暮らしていた）日本語母語話者であるが、ほかに日本でドイツ語を学ぶ外国人（非日本語母語話者）なども含めて考えてよいであろう。
- 2) 本研究は科学研究費助成事業（基盤研究(C)25370478）の助成を受けている。
- 3) Duden (2013: 134) による。そこでは平均的な母語話者の能動的 (aktiv) な語彙は12,000～16,000語、受動的 (passiv) な語彙は少なくとも50,000語であるとされている。
- 4) この図では示されないが、これ以外に現代性も問題となる。ただし基本語彙では通常現代語のみが問題となるのでここでは立ち入らない。

らの語が各領域に収まるかは不明である。⁵⁾ 本稿では、「高い頻度」を示し、「広い使用域」に出現する語、すなわち A に位置する語をまず基本語彙の中核として設定する。次に使用域の広さを重視する。つまり、A の次は C ではなく B の部分が優先される。そして最終的に日本人にとっての頻度という点を考慮する。日本人にとって高い頻度で用いられる語を、母語話者の使用状況と照らし合わせて調整しようというものである。その場合も使用域の広さは重視される。理論上は、下の図の B（ないし B~D の境界辺り）に位置する語を A（ないし A~C の境界辺り）に引き付けて調整するということになる。⁶⁾

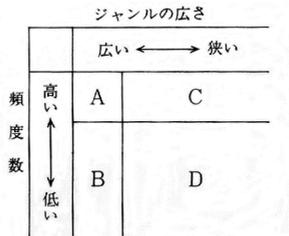


図1 語彙のあり方 (千野1986: 59)

1.2 先行研究

基本語彙を選定しようという試みはもちろん新しいものではない。日本国内に限っても、仁科ほか (1971), 朝倉ほか (1982), 重藤 (1990), 朝倉ほか (1994), 石川/佐藤 (1996), 岩崎 (2012) などで、その試みやさまざまなアプローチについて知ることができる。また出版市場では数多くの単語集が出回っている。このような状況の中、現時点であらためて基本語彙について検討しようとする最も重要な理由は、大規模コーパスの登場である。IDS の DeReKo など、複数の大規模コーパスが利用可能となり、それに基づく語

- 5) 市販の単語集では、テキストカバー率を根拠に2,000~4,000語を記載していることが多い。しかしながら、たとえ80~90%程度のカバー率があったとしても、実際の有用性は甚だ疑問である (Ickler 1984: 21f.)。また、これまでの検証によると、単語集で謳われているカバー率はいささか高く見積もられ過ぎているようである (朝倉ほか 1994, 黒田 2014 など)。本稿の調査結果 (後述) も参照。
- 6) 例えば、Japan, japanisch といった語は頻度上1000~2000位レベルに位置するが、日本人学習者のための基本語彙を考える場合は、もっと上位になるようランクの調整が必要であろうということである。一方、日本に特徴的な語であっても、特殊なもの (例えば Sake, Reiswein など) は——ドイツ語圏の各種コーパスに幅広く出現しないという限りにおいて——本稿で言うところの基本語彙には含まれない。それらの語は別途まとめられ、学習されることになる。本稿の第4節も参照。

彙リストが参照できるようになった（在間 2009, 恒川 2014 を参照）。その有用性を検証することはドイツ語教育に携わる者の急務である。

Kaeding (1898) による1千万語を越える、手作業としては驚くべき規模の語彙調査を例外とすれば、従来の基本語彙選定は、比較的小規模の調査に基づくか、選定者による特定の観点に基づくもの、あるいはその両者を併用したものであったと言ってよいであろう（Kühn 1990, 江口 1996 を参照）。これ以外に、既存の単語集・辞書における重要語の重複状況を調べ——いわばその最大公約数をとって——基本語彙を選定するという試みもある（例えば重藤 1990, 石川／佐藤 1996 など。今道 2009 も参照）。

特筆すべきこととして、かつて独文学会内でも語彙調査が行われている（仁科ほか 1971）。報告によると、中国・四国の会員を中心に、総語数約117,000語、見出し語（レマ）にして約17,000語の、手作業による語彙調査が3年以上をかけて行われたとのことである。しかしながら、それらの成果が——例えばネット上などで——容易に参照できる状況にあるかと言えば、必ずしもそうっていないというのが現状であろう。⁷⁾ なお、出版されている頻度辞典として Jones/Tschirner (2006) がある。Tschirner (2008) で示された追加語を含めると全体として約5,000語の語彙リストとなるが、大藪 (2014) では、このリストを他の頻度語彙リストと比較し、かなりの程度ずれが認められることを確認している。本稿では、さらに教育的観点からの語彙リストも含め調査を進める。

1.3 比較のための準備

語彙リストを比較する際に、極めて重要であるが、しかし同時に大変困難な作業は、見出し語（レマ）の基準をそろえることである。極端な例として、Start Deutsch (SD1) の語彙リストには gestorben という見出し語がある（例文は Meine Frau ist gestern gestorben.）。通常の語彙リストであれば、見出し語として sterben を立てるところであろう。SD1に sterben がなく、他のリストに gestorben がないという状況を仮に「差分2」として記述したとしても、表層的な比較にしかなり得ない。変化形以外にも、派生形、複合語、固有名詞、外来語、略語、記号などの取り扱いが問題となる。

見出し語の立て方は、そもそも基本語彙選定の際にも問題となるものである（在間 2012）。本稿では、大藪 (2014: 53f.) で立てた基準に依拠する。

7) 石川／佐藤による次のコメントも参照。「ドイツ語の授業における語彙の取り扱いに関しては、これまでも多くのドイツ語教師が様々な試みを実行し、またかなりの成果をあげている場合も多いと思われるが、残念なことにそれらの報告を目にする機会は少ない」（石川／佐藤編 1996: はじめに）。

詳細は繰り返さないが、1点だけ触れておくと、変化形・派生形などが問題となる場合は、既存の独独辞典を参照し決定する。具体的には、*Duden Universalwörterbuch*, *Langenscheidt Großwörterbuch Deutsch als Fremdsprache*, *PONS Großwörterbuch Deutsch als Fremdsprache* の3冊を参照し、いずれの辞書においても見出し語として扱われている場合に限り見出し語として取り上げる。

2. 語彙リスト

以下、語彙リストの調査を進めていくが、使用する語彙リストは、大きく、頻度に基づいて作成された基本語彙リスト、教育的観点から作成された基本語彙リスト、並びに独自調査に基づく語彙リストの3種類である。なお、個別の語彙リストについては以下略号で示す。本稿末尾の略号説明を参照のこと。

2.1 頻度語彙リスト

大藪（2014）で作成したリストを利用する。これは各種コーパスに基づく4つの頻度語彙リスト DRW, DWC, J&T, DtW に基づき5,000語を選定したものである（各リストの特徴については山田 2009 も参照）。なお、5,000語という数は、とりわけ英語を対象とした研究で語彙学習の一つの目安とされているもの（例えば Nation 2001）を出発点として仮に採用したものである。次の5段階（下位区分を含めると6段階）にレベル分けをしてある。この5,000語のリストを以下「リストD」と呼ぶ。

表1 「リストD」のレベル設定

レベル	語数	語彙レベル	頻度順位
レベル1A	500語	500語レベル	1～500位
レベル1B	500語	1,000語レベル	500～1,000位
レベル2	1,000語	2,000語レベル	1,001～2,000位
レベル3	1,000語	3,000語レベル	2,001～3,000位
レベル4	1,000語	4,000語レベル	3,001～4,000位
レベル5	1,000語	5,000語レベル	4,001～5,000位

この「リストD」のもとになった各個別リスト間の重複状況は次の通りである。○はリストに含まれていることを、×は含まれていないことを示す。

表2 「リストD」詳細

DRW	DWC	J&T	DtW	語数	リスト別語数
○	○	○	○	3,837 (4リスト重複: 3,837語)	DRW 4,898語
○	○	○	×	450	DWC 4,952語
○	○	×	○	562 (3リスト重複: 1,125語)	J&T 4,409語
○	×	○	○	40	DtW 4,539語
×	○	○	○	73	
○	○	×	×	5	
○	×	○	×	1	
○	×	×	○	3 (2リスト重複: 37語)	
×	○	○	×	4	
×	○	×	○	21	
×	×	○	○	3	
×	×	○	×	1 (1リストのみ: 1語)	
計				5,000	

原則として3リスト以上に共通する語を採用するという方針を取ったが、上の表から分かる通り、38語については2リストまたは1リストにしか挙がっていない。これらは個別の事情を配慮して採用を決めた語である。例えば次のような場合がある。

まずDRWでは略語 (bzw. など) やハイフンでつながれた複合語 (E-Mail など) は見出し語化されていない。J&Tでは固有名詞は本文見出し語に含まれず、巻末で100位までの語が確認できるのみである。またDtWはレマ化を行っていない語形リストであるため、分離動詞は上位に出現しにくい。ほかにリストごとに異なる語形に見出し語化されているケースがあったり、レマ化のミスではないかと思われるケースもある。⁸⁾ 参考までに、1リストにしか挙がっていないにもかかわらず採用した語は *mithilfe* である (J&Tのみ)。これは旧正書法で *mit Hilfe* と分かち書きされていたものが、新正書法で一語書き・分かち書きともに認められることになった例である。最新のDuden (2013) では *mithilfe* を推奨形としているので、本リストでもこの形を採用している。

以上のようなケースでは、他の語形の頻度を見ることによってレマ化の過程を推測できる場合もある。例えば、J&Tにおいて *ausgeglichen* は見出し語にないが、*ausgleichen* の順位が他のリストと比べて高いため、両者がまとめて処理されたのだと考えられる。同様に、J&Tでは *auseinandersetzen* という見出し語がないが、*auseinander* の順位が他リストより高いことなどから、

8) DWCは特に語末の文字を落としてレマ化する傾向がある。例えば、*-e* が落ちたのではないかと思われる例として *Eck, Sekund, Spitz, Stadtwerk* など。*-n* が落ちたのではないかと思われる例として *Abkomme, Esse, Friede, Kaste, Lade, Niedersachsen, Streife* など。

auseinander setzen という1996年～2006年の過渡期の正書法で処理されたのではないかと推測される。以上、見出し語の立て方の難しさが見取れると同時に、複数の語彙リストが利用可能だからと言って、単純な統計的処理によって合成を試みるといった方法を取ることに問題があることが分かる。精度の高い語彙リスト作成のためには、現在のところ慎重な手作業が不可欠である。

2.2 教育語彙リスト

SD1, SD2, ZD（それぞれヨーロッパ言語共通参照枠のA1, A2, B1レベル相当）並びに市販の単語集BGを使用する。

まずSD/ZDは、それぞれ Wortgruppenliste と Alphabetische Liste から成るが、両者をまとめて扱う。SD1については、2009年の版と2011年の版で若干の相違があるが、本稿ではどちらかの版に記載されていればすべてSD1として扱った。また、SDにあってZDにない語が少なからずあるが、それらも単純にSDレベルの語として処理した。本稿の基準に従って見出し語の調整を行った結果、SD1レベルが797語、SD2レベルが577語、ZDレベルが1,403語の見出し語数となった。合わせて2,777語である。

BGは4,000の見出し語が、1-2000と2001-4000という2つのレベルに分けて記載されている。⁹⁾ 最終版の出版は1991年であるが、奥付に2006年の正書法改革に対応しているとあり、1991年以降に訂正が施されたと思われる箇所もある（Mark, Pfennig から Euro, Cent など¹⁰⁾）。本稿の基準で整理し直すと、総見出し語数は3,638となった。そのうち1-2000レベルが2,174語、2001-4000レベルが1,464語である。¹¹⁾

SD/ZD および BG の重複状況は次の通りである。合計4,238語のこのリストを、以下「リスト DaF」と呼ぶことにする。

表3 「リスト DaF」内訳

SD/ZD	BG	語数	リスト別語数	
○	○	2,177	SD/ZD	2,777語
○	×	600	BG	3,638語
×	○	1,461		
		計		
				4,238

- 9) 同じ語形の語が複数のページに記載されていたり、名詞の男性形に女性形が併記されたりしているが、それらをすべて数え上げると4,000語になる。なお、巻末のIndexと本体の間かなりの不一致が見つかるので注意が必要である。
- 10) ただし、Internet, E-Mail, Handyなど、現在では重要語と考えられるいくつかの語は含まれていない。
- 11) 一部レベルの記載がない語もあるが、前後の状況などから1-2000レベル相当であると判断した。

SD/ZD と BG に共通するのは2,177語である。ところで、教育語彙リスト間のずれを指摘した興味深い報告として Kühn (1990) がある。Kühn が独自に6つの語彙リスト——中にはやや古いものも含まれる——の R の項を取り上げて比較したところ、全体として221の語があったにもかかわらず、6リストすべてに共通していたのはわずか4語であったとのことである (Kühn 1990: 1355)。ちなみに SD/ZD および BG において R で始まる語の総数は148語、そのうち両リストに共通しているのは73語である。ついでながら、先の DRW, DWC, J&T, DtW と合わせるとちょうど6リストになるので、R で始まる語の総数を調べてみたところ、偶然221語であった。そのうち6リストすべてに共通する語は51語である。どうやら Kühn の挙げている例はやや極端に過ぎるようである。

3. 比較

3.1 頻度語彙リストと教育語彙リスト

上述の「リスト D」と「リスト DaF」について、まず全体の重複状況は次の通りである。

表4 「リストD」・「リスト DaF」比較

リスト D	リスト DaF	語数	リスト別語数	
○	○	2,949	リスト D	5,000語
○	×	2,051	リスト DaF	4,238語
×	○	1,289		
		計	6,289	

「リスト DaF」の4,238語のうち、1,289語が「リスト D」(5,000語)に含まれていないことが分かる。そこで次に、これらの語の頻度をもう少し正確に把握するため、「リスト D」を構成する各個別リスト全体を再度調査してみた。その結果、691語がもとの個別リストの少なくともいずれかに含まれていることが分かった。この691語を「リスト D」に含めて見直すと、重複状況は次のようになる。

表5 「リストD(拡大版)」・「リスト DaF」比較 (1)

リストD(拡大)	リストDaF	語数	リスト別語数	
○	○	3,640 (+691)	リストD(拡大)	5,691語 (+691語)
○	×	2,051	リスト DaF	4,238語
×	○	598 (-691)		
		計	6,289	

以下、個別リストの重複状況について、興味深いと思われるパターンを見ていく。まず次は6つの個別リストすべてに共通している語である。全部で1,631語ある。さらに各リストの上位語に限定すると、その数は260語となる。

表6 「リストD (拡大版)」・「リスト DaF」比較 (2)

DRW	DWC	J&T	DtW	SD/ZD	BG	語数
○	○	○	○	○	○	1,631

レベル (上位のみ)	例	語数
リスト D: 1A リスト DaF: SD1 & 1-2000	ab, aber, alle, allein, als, also, alt, an, anbieten, andere, ... zu, zusammen, zwei, zweit, zwischen	260

これらは基本語彙の中核を成すと考えてよいであろう。次の表7は、「リスト D」内のすべての個別リストに共通しているにもかかわらず、「リスト DaF」には存在しない語の例である。大ざっぱに、「母語話者にとっては高頻度だが、ドイツ語教育ではあまり重要視されていない語」と考えてよいであろう。全部で1,272語ある。頻度レベル別に例も示す。

表7 「リストD (拡大版)」・「リスト DaF」比較 (3)

DRW	DWC	J&T	DtW	SD/ZD	BG	語数
○	○	○	○	×	×	1,272

リストD レベル	例	語数
レベル1A	Angabe, Bereich, Berlin, darstellen, ... wirken	13
レベル1B	AG ¹²⁾ , Aktion, Anspruch, aufgrund, ... zuvor	57
レベル2	Abschluss, absehen, akzeptieren, allzu, ... zurückziehen	197
レベル3	Abbau, abbauen, Abbildung, abhalten, ... zustimmen	381
レベル4	abbrechen, Abenteuer, ablösen, absurd, ... zwangsläufig	393
レベル5	abbilden, abreißen, absagen, absehbar, ... Zyklus	231

続いて次の表8で示されているのは、この逆のパターン、つまり「リスト D (拡大版)」にはまったく記載がないにもかかわらず、「リスト DaF」ではいずれの個別リストにも記載のある語である。「母語話者にとっては低頻度だが、ドイツ語教育では重要視されている語」であると考えられよう。これらは本稿の目的と照らして注目すべき語である。全部で126語ある。ここでもレベル別に例を示す。

12) DRW では本来略語は取り上げない方針であるが、Ag という見出し語が含まれていた。「銀」(Argentum) を指すとは考えにくいので AG として処理してある。

表8 「リストD (拡大版)」・「リスト DaF」比較 (4)

DRW	DWC	J&T	DtW	SD/ZD	BG	語数
×	×	×	×	○	○	126

リスト DaF レベル 例	語数
SD1 & 1-2000 anmachen, Birne, duschen, einundzwanzig, ... Wochentag	10
SD1 & 2001-4000 abfliegen, Abflug, Absender, Bleistift, ... Schinken	12
SD2 & 1-2000 beeilen, Bluse, bluten, erkälten, ... zumachen	19
SD2 & 2001-4000 Bohne, Brieftasche, Briefumschlag, Feuerzeug, ... Zitrone	12
ZD & 1-2000 abmachen, abtrocknen, abwärts, abwesend, ... WC	33
ZD & 2001-4000 abmelden, Badewanne, Briefträger, Camping, ... Zahnpasta	40

Tschirner (2005) は、頻度語彙リストである Jones/Tschirner (2006) と教育語彙リストの BG (両者とも約4,000語) の重複状況を調査し、BG の語のうち約40% が Jones/Tschirner に含まれていなかったと報告している。例として挙げられている語のうち、wirken, zuvor などは上の表7 の例でも確認できる。BG が頻度を反映していないことに対し Tschirner は批判的である。だがその批判はどの程度妥当性のあるものであろうか。「頻度語彙」を示す上の表7 の例は、直感的に文語的な傾向が強いように感じられる。それとは対照的に、「学習語彙」を示す表8 の例は、日常的に身近な事物を指す語を多く含んでいるという印象を受ける。興味深いことに、仁科 (1971 : 532) も、独自に作成した頻度語彙リストについて、「日常身近かな事・物の出現回数 は高くない」と同様の感想を述べている。

岩崎 (2012) は、基本語彙選定において頻度順アプローチを取ることの問題点について論じているが、とりわけコーパスの均衡性の問題、そして日本における教育という特殊事情のもたらす問題を指摘している。岩崎も触れているが、特に DRW や DtW では新聞などで用いられる書き言葉の比重が高い (ただし山田 2009 によると、DWC と J&T はそれに対し話し言葉的な傾向を示す)。確かに Jones/Tschirner (2006) は均衡コーパスに基づいた語彙リストであるが、しかしながら、何ををもって「均衡」と見なすかについては、そもそも決定的な答えはない。基本語彙リストを学習に応用しようとした場合、頻度情報に加え、「学習語彙」の情報も取り込むことは、少なくとも現時点においては有用であると考えられる。なお、今回は紙幅の関係で、リスト全体の重複状況のみを取り上げたが、細かく見ればレベル (ランク) の相違もある。そのような相違も踏まえて、今後、基本語彙の調整を考えていく必要がある。

3.2 独自調査

本節では、日本人にとっての基本語彙抽出を目的として行った独自調査について報告する。ここで作成したリストは、これまでのものとは異なり、基本語彙以外の語も含むものである。なお、このリストについてはすでに大藪(2014)で一部報告した。調査対象としたのは、小島泰(監)『ドイツ人が日本人によく聞く100の質問(改訂版)』(以下100F)と福本義憲『ドイツ語会話110番』(以下K110)の2冊の書籍である。前者は日本社会・文化を説明した文章(書き言葉的)、後者は日本人を登場人物とした日常会話集(話し言葉的)であり、前者は100の、後者は110のテーマから成る。両者ともに文脈のある、ある程度まとまったテキストを構成しているため、調査資料として都合がよいと考えた。

両者を合わせた総語数は21,218、見出し語(レマ)数は3,654である。¹³⁾ 実際には相当数の語が1回しか出現しない(具体的には21,218語中の2,138語)。本稿では1回しか出現しない語(粗頻度1の語)は、リストから除外するものとした。さらに、人名、アラビア数字、単位記号など、本稿の基準では見出し語の資格のない17語をのぞくと、結果として1,499語の見出し語が残った。調査した二つのテキストはそれぞれの総語数が異なるので、粗頻度ではなく相対頻度で両者の平均を出し、全体の順位を出しておく。1,499語から成るこの語彙リストを、以下「リストJ」と呼ぶことにする。

100F	K100	語数	テキスト別採用語数	
○	○	414	100F	1,413語
○	×	999	K100	500語
×	○	86		
		計	1,499	

まず、「リストD」(5,000語)と「リストDaF」(4,238語)が、上記2冊の本を実際のところどの程度カバーするのか(全体の述べ語数に対するカバー率)を見ておきたい。

テキスト	総語数	カバー語数	カバー率
100F	17,077	14,547	85.2%
K110	4,141	3,858	93.2%
合計	21,218	18,405	86.7%

テキスト	総語数	カバー語数	カバー率
100F	17,077	14,224	83.3%
K110	4,141	3,896	94.1%
合計	21,218	18,120	85.4%

13) ミスを訂正したため大藪(2014)の数字と異なっている。

カバー率に大きな差はないことが分かる（もっとも、上述の通り全体の10%程度の語の頻度は1であるため、たとえ数パーセントであっても違いの持つ意味合いは小さくない。逆に言うと、カバー率を1%上げるのは、90%前後からは非常に難しくなるということである）。『会話110番』（K110）のカバー率に関しては、むしろ「リスト DaF」の方が「成績」が良い。ただし、これは学習参考書というテキストの性質によるものであろう。一方で『100の質問』（100F）のカバー率が両者ともに意外に低いという印象があるかもしれない。しかし、こちらの方は、日本特有の固有名詞を多く含むこと、さらに日本語のローマ字表記（*geisha*, *ryokan* など）が多用されていることも影響していると考えられる。

では「リストD」と「リストJ」を比較してみる。

表12 「リストD」・「リストJ」比較

リストD	リストJ	語数	リスト別語数	
○	○	1,112	リストD	5,000語
×	○	387	リストJ	1,499語
		計	1,499	

「リストJ」の見出し語1,499語のうち、「リストD」に含まれるのは1,112語である。つまり、387語が「リストD」に含まれていない。そこで、ここでも「リストD」のもとになっている各個別リストを再調査してみた。その結果、117語がもとの個別リストの少なくともいずれかに含まれていることが確認できた。

表13 「リストD（拡大版）」・「リストJ」比較（1）

リストD(拡大)	リストJ	語数	リスト別語数	
○	○	1,229 (+117)	リストD(拡大)	5,117語 (+117語)
×	○	270 (-117)	リストJ	1,499語
		計	1,499	

以下、前節と同様、重複状況が特徴的な組み合わせを取り上げてみる。次はすべての個別リストに共通する語である。上記1,229語のうち1,007語が該当する。

表14 「リストD」（拡大版）・「リストJ」比較（2）

DRW	DWC	J&T	DtW	リストJ	語数
○	○	○	○	○	1,007

例：ab, Abend, aber, abnehmen, Abschluss, ... zwei, zwischen

これらは頻度も高く、使用域も広い語であると考えられる。なお、「リストD」および「リストJ」における各語のレベル／ランクを比べてみると、概ね平行していることが分かる。興味深いのは、そのような中でレベル／ランクに不一致の見られる事例である。「リストJ」の上位の方からいくつか例を挙げると次のようなものがある。いずれも「リストD」の頻度レベルに比して「リストJ」で順位の高い事例である。

表15 「リストD (拡大版)」・「リストJ」比較 (3)

見出し語	リストDレベル	リストJ順位
Japan	2	27
bitte	2	28
japanisch	2	48
Dank	2	60
ach	2	62
Entschuldigung	5	105

調査対象としたテキストの性質からして、日本社会・文化に関わる語彙および日常会話の語彙が抽出されると予想されるが、実際にそのような語が挙がっているようである。

次に示すのは「リストD」の個別リストにまったく記載のない語である。270語ある。

表16 「リストD (拡大版)」・「リストJ」比較 (4)

DRW	DWC	J&T	DtW	J	語数
×	×	×	×	○	270

例：Abschlusszeremonie, Abwanderung, Ainu, Ainu-Sprache, *akibakei*, ... Zweikammersystem

これらは頻度が低いか、使用域が極めて限られているか、あるいはその両方のタイプの語であると見なすことができる。このタイプの語は差し当たり本稿の関心の対象ではない。ただし、先の「リストDaF」との関係で捉え直すと、興味深い分布を示す語もわずかながら存在する。

表17 「リストD (拡大版)」・「リストDaF」・「リストJ」比較

DRW	DWC	J&T	DtW	SD/ZD	BG	リストJ	語数
×	×	×	×	○	○	○	8

全該当例：beelen, duschen, einpacken, frieren, frühstücken, heizen, mittags, Speisekarte

頻度語彙リストには出現しないにもかかわらず、教育語彙リスト並びに独自調査リストの両方に含まれる語である。考慮が必要なパターンであろう。

4. おわりに

以上、語彙リストの比較を同種の語彙リスト間および異なる種類の語彙リスト間で行った。とりわけ着目したのは、頻度語彙リストと教育語彙リスト間の異同、並びに独自調査による語彙リストで浮かび上がってくる語である。頻度語彙リストの5,000語のうち、教育語彙リストと共通するのは約3,000語であった。両者の語彙はまったく一致しないとまでは言い切れないが、一方で無視することもできない程度の相違があると言えよう。今後は、独自調査の結果も踏まえ、各リストの結果を合成していくという作業に進むことになる。ただし、2.1で見たように、見出し語化の基準や精度の問題が大きいため、単純な統計学的手法を取ることは今のところ困難である。

最後に、語彙教育と基本語彙の関係について一言だけ触れておきたい。基本語彙というのは、その特性からして多くのテキストに共通の、特殊でない語の集合である。一方、具体的な個々のテキストは、ふつう何か特別なことを伝えるために書かれたり、発話されたりするものである。Ickler (1984: 21) の言葉を借りるならば、「テキストの特殊でない90%は、いわば、内容に富んだ10%を際立たせる背景のようなもの」というふうにも考えることもできる。テキストの正確な理解にとって、基本的でない語の理解が重要なのは言うまでもない。ジャンルごとの高頻度語（冒頭の図で言うとCの部分）の取り扱いについても、今後さらに積極的に考えていく必要があるだろう。

略号（語彙リスト）

100F = 『ドイツ人が日本人によく聞く100の質問』の語彙リスト。独自調査による。

BG = *Langenscheidt Basic German Vocabulary*. Berlin: Langenscheidt. 1991 (2006年新正書法対応版)。

DRW = DeReWo (IDSによる、大規模コーパスDeReKoに基づく語彙リスト)。derewo-v-30000g-2007-12-31-0.1 (Grundformliste) のうち上位1万語を参照。

DtW = *Deutscher Wortschatz* (Leipzig 大学情報学研究所のプロジェクトによる、主に新聞コーパスに基づく語彙リスト)。1万語の語形リストを独自にレマ化して参照。

DWC = deWaC (Trento 大学 Baroni 氏ほかによる、de ドメインを利用した Web コーパスに基づく語彙リスト)。deWaC unigrams (lemmas) のうち数字・記号などをのぞいた上位1万語を参照。

K110 = 『ドイツ語会話110番』の語彙リスト。独自調査による。

J&T = Jones/Tschirner (2006) 並びに Tschirner (2008) の計5,000語を参照。

SD1 = Goethe Zertifikat A1. Start Deutsch 1. Prüfungszeile, Testbeschreibung. 2. aktualisierte Aufl. Goethe-Institut. 2011 並びに下記の SD2 も参照。

SD2 = Start Deutsch. A1, A2. Prüfungszeile, Testbeschreibung. 2. Aufl. TELC. 2009 を参照。

ZD = Zertifikat Deutsch. Lernziele und Testformat. TELC. 1999 を参照。

参考文献

- 朝倉巧ほか (1982) : 「Mindestwortschatz 策定の試み (1)」『ドイツ語教育部会会報』21, 51-58.
- 朝倉巧ほか (1994) : 「基礎語彙策定の試み」『ドイツ語教育部会会報』46, 5-14.
- 千野栄一 (1986) : 『外国語上達法』岩波書店.
- Duden (2013): *Die deutsche Rechtschreibung*. 26. Aufl. Mannheim: Dudenverlag.
- 江口豊 (1996) : 「ドイツ語圏の辞書・辞書学と基礎語彙」石川／佐藤 (編), 1-31.
- Ickler, Theodor (1984): *Deutsch als Fremdsprache. Eine Einführung in das Studium*. Tübingen: Niemeyer.
- 今道晴彦 (2009) : 「ドイツ語教育における重要語彙選定の現状と課題」『ドイツ語教育』14, 40-47.
- 石川克知／佐藤俊一 (1996) : 「基礎語彙リスト」石川／佐藤 (編), 57-101.
- 石川克知／佐藤俊一 (編) (1996) : 『ドイツ語基礎語彙研究』(北海道大学言語文化学部研究報告書8).
- 岩崎克己 (2012) : 「日本の DaF におけるドイツ語基本語彙へのアプローチ」岡村ほか (編), 45-66.
- Jones, Randall L./Tschirner, Erwin (2006): *A Frequency Dictionary of German*. London: Routledge.
- Kaeding F. W. (1898): *Häufigkeitwörterbuch der deutschen Sprache*. Steglitz: Selbstverlag des Herausgebers.
- 黒田廉 (2014) : 「Web ページの内容理解と辞書の収録語彙」恒川／大藪 (編), 27-36.
- Kühn, Peter (1990): ‚Das Grundwortschatzwörterbuch‘. In: Hausmann, Franz Josef (Hg.): *Wörterbücher*. Berlin: de Gruyter, 1353-1364.
- Nation, I. S. P. (2001): *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- 仁科武光ほか (1971) : 「教授用重要語彙の統計的選定」日本独文学会ドイツ語学委員会 (編) : 『ドイツ語の基本的諸問題』南江堂, 231-534.
- 岡村三郎ほか (編) (2012) : 『ドイツ語基本語彙——辞書学的, 外国語教授法的な観点から』(日本独文学会研究叢書88).
- 大藪正彦 (2014) : 「基本語彙と頻度——実践と課題」恒川／大藪 (編), 49-64.
- 重藤実 (1990) : 「ドイツ語教育と基礎語彙」『ドイツ語教育部会会報』37, 18-22.
- Tschirner, Erwin (2005): ‚Korpora, Häufigkeiten, Wortschatzerwerb‘. In: Heine, Antje u.a. (Hg.): *Deutsch als Fremdsprache*. München: Iudicium, 133-149.
- Tschirner, Erwin (2008): ‚Das professionelle Wortschatzminimum im Deutschen als Fremdsprache‘. *Deutsch als Fremdsprache* 4/2008, 195-208.
- 恒川元行 (2014) : 「コーパス利用に基づくドイツ語研究——幅広いデータ収集と頻度から見直す」恒川／大藪 (編), 2-11.
- 恒川元行／大藪正彦 (編) (2014) : 『コーパス利用に基づくドイツ語研究——幅広いデータ収集と頻度から見直す』(日本独文学会研究叢書98).
- 山田善久 (2009) : 「語彙使用頻度調査は今どこまで可能になっているか」在問進／時田伊津子 (編) 『大規模コーパスを用いたドイツ語研究』(日本独文学会研究叢書61), 23-35.
- 在問進 (2009) : 「ドイツ語コーパスの現状」『国文学——解釈と鑑賞』1月号, 150-157.
- 在問進 (2012) : 「頻度に基づく「基本語彙」リスト作成の問題点と展望」岡村ほか (編), 3-12.

Vergleiche deutscher Grundwortschatzlisten

Masahiko OZONO

Die neuere Entwicklung der Sprachkorpora und Computerprogramme hat die Erstellung von Wortlisten aus Milliarden Wörtern (Wortformen) ermöglicht. Um die häufigkeitsbasierten Wortlisten auf den Deutschunterricht in Japan anzuwenden, müssen aber zumindest noch zwei Fragen geklärt werden: Erstens stimmen die häufigen Wörter nicht immer mit denen überein, die für Lerner Vorrang haben. Zweitens deckt sich die Häufigkeit für deutsche Muttersprachler nicht unbedingt mit der für japanische Deutschlerner, die in Japan leben oder gelebt haben. Das Ziel des vorliegenden Beitrags besteht darin, zunächst einmal die Diskrepanz zwischen den für deutsche Muttersprachler häufigen Wörtern und den für japanische Deutschlerner wichtigen Wörtern auf konkrete Wörter hin zu prüfen.

Hierzu werden Vergleiche zwischen drei Typen von Grundwortschatzlisten angestellt: (1) Frequenzlisten, die auf verschiedenen Korpora beruhen. Die aus einzelnen Wortlisten zusammengefaßte Wortliste mit den häufigsten 5 000 Wörtern wird in dem Beitrag „Liste D“ genannt. (2) Wortlisten aus dem DaF-Bereich, darunter z.B. die Wortliste für die Sprachprüfung „Start Deutsch“ oder „Zertifikat Deutsch“. Die zusammengefaßte Wortliste mit ca. 4 200 Lernwörtern wird „Liste DaF“ genannt. (3) Zur Ermittlung der wichtigen Wörter für japanische Lerner wurde eine eigene Wortschatzuntersuchung anhand von zwei Büchern (*Hundert Fragen über Japan* und *Deutsche Konversation 110*) durchgeführt. Von insgesamt ca. 21 200 Wörtern (Wortformen) wurde schließlich eine Wortliste mit ca. 1 500 häufigen Wörtern (Grundformen) erstellt. Die Liste wird „Liste J“ genannt.

Um das Ergebnis vorwegzunehmen: Von 5 000 Wörtern der Liste D decken sich etwa 3 000 mit denen der Liste DaF. Zwar kann man hier nicht von einer kompletten Nichtübereinstimmung sprechen, aber es besteht doch ein signifikanter Unterschied. Ein Vergleich zwischen Liste D und J offenbart erwartungsgemäß, daß bestimmte Wortgruppen in Liste J einen höheren Rang haben: Bei der einen Wortgruppe handelt es sich generell um das Land Japan (z.B. *Japan, japanisch*), bei der anderen um Wörter, die zur alltäglichen Konversation gehören (z.B. *bitte, [vielen] Dank*). Die nächste Aufgabe ist nun, die Wortlisten aufgrund der Erkenntnisse zu einer Grundwortschatzliste für japanische Lerner zusammenzuführen.